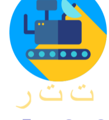
تويشيات فالتعلم الآلي - الانحدار الخطي

Redouane Lguensat
fb.com/AppliedMathsInDarija

Simple introduction to Machine Learning in Moroccan Arabic

March 2020



السلام عليكم

1. هاد التقديم محاولة متواضعة باش نوصل بعض الافكار على التعلم الآلي Machine Learning بشكل مبسط و بالدارجة المغربية
2. عانحاول يكون هاذ التقديم سهل للفهم و لكن خاص اللي كايقرا ولا كاتقرا هادشي يكون عندهم شوية مع الرياضيات و المعلوميات
3. الناس اللي باغين يقراو التعلم الآلي خاصهم يتعمقو اكثر فالدروس المتوفرة على الانترنت، هادشي اللي كاندير عا ديك الدغمة الاولى "لمحاربة الأمية" فهاد الدومين



مقدمة

تخيل معايا عطيتك هاد الأزواج ديال البيانات X وY

X	Y
3	6
7	14
9.5	19

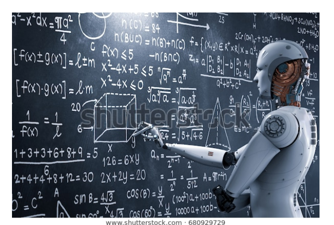
ايلا قلت ليك را كاينة علاقة بيناتهم و سولتك شنو قيمة Y ايلا كان X هو 5 ؟


www.shutterstock.com · 360943928



مقدمة

ايلا فكرتي شوية عا يبان ليك را الجواب هو Y=10 حيث النموذج الرياضياتي $Y = 2X$ صالح لذوك البيانات



التعلم الآلي (ولا النمذجة الإحصائية Statistical Modeling بشكل عام) مبنية على هاذ القضية اللي درتي دابا: الخوارزمية كاتشوف البيانات و كاتقلب على النموذج الرياضياتي الأحسن اللي كايقدر يكون صالح للبيانات

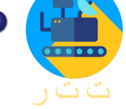
شناهو الانحدار الخطي؟

فهاذ الحلقة عادي ندويو على اسهل نموذج رياضياتي فالسوق، اللي هو أول حاجة كايبداو بيها الناس اللي كايبغيو يدخلو للدومين ديال التعلم الآلي: هاذ النموذج هو الانحدار الخطي Linear Regression

1. النموذج اللي كانفترضو انه مزيان للبيانات كايكون على شكل دالة تآلفية

$$Y = \alpha X + \beta \quad (1)$$

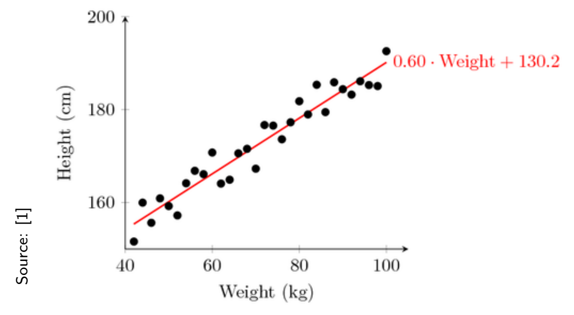
2. X و Y يقدرو يكون متجهات عامرين بعدة أعداد. و لكن هنا عانبقاو فأسهل مثال اللي هو ملي كايكون البعد Dimension ديالهم هو 1، يعني فيهم قيمة وحيدة اللي كاتغير فالبيانات (بحال المثال اللي فالمقدمة)
3. فهاذ الحالة α و β عددان حقيقيان مستقلان عن X





الانحدار الخطي البسيط Simple Linear Regression

1. ايلا رسمنا النموذج اللي دوينا عليه سابقا ، من الواضح انه عا يكون مبيان هذه الدالة مستقيما معامله الموجه هو α و β هو أرتوبه عند الصفر
2. هذا مثال ديال بيانات تاع الوزن و الطول ديال 30 واحد، ايلا رسمناهم فمبيان و لقينا احسن انحدار خطي، عانلقاو ان $\alpha = 0.06$ و $\beta = 130.2$ و الدالة التآلفية كاتبان قريبة بزاف من البيانات و كاتأكد أن النموذج مابيهش





كيفاش درنا تا لقينا هاذ الانحدار الخطي البسيط؟

1. باش تلقى الانحدار الخطي البسيط اللي يجي مع البيانات ديالك، خاصك تلقى α و β
2. ايلا شفتي المثال السابق عا تلاحظ ان الدالة ما كاتدوزش ديريكت من كاع البيانات، كاين واحد التيساع بين البيانات و الدالة، وهاذ التيساع هو الخطأ Error اللي دارتو الدالة فذيك القيمة. لنفترض أن فكل نقطة i درنا خطأ ϵ_i يعني: $Y_i = \alpha X_i + \beta + \epsilon_i$
3. ايلا بغينا الانحدار الخطي البسيط ديالنا يكون مزيان، خاص المجموع ديال الأخطاء فكاع البيانات يكون صغير.
4. باستعمال واحد الطريقة سميتها طريقة المربعات الصغرى أو الدنيا Least Squares عا يكون الهدف هو نلقاو $\hat{\alpha}$ و $\hat{\beta}$ باش نصغرو الجمع ديال المربعات تاع الأخطاء

$$\hat{\alpha}, \hat{\beta} = \min_{\alpha,\beta} \sum_{i=1}^{n} \hat{\varepsilon}_i^2 = \min_{\alpha,\beta} \sum_{i=1}^{n} (Y_i - \beta - \alpha X_i)^2$$



الحل

(تمرين) دير الاشتقاق و لقا القيم الدنيا اللي كاتحل المشكل الرياضياتي.
الحل هو هذا:

$$\hat{\alpha} = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2} \tag{2}$$

$$\hat{\beta} = \bar{Y} - \hat{\alpha}\bar{X} \tag{3}$$

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \quad \bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \tag{4}$$

ايلا عندك مع المعلوميات كاين بزاف ديال البرامج اللي كايحسبو ليك هاذ الحل، مثلا ف Python شوف التالي



شوية ديال الكود

```
from sklearn.linear_model import LinearRegression

regr = linear_model.LinearRegression()
regr.fit(X, Y)
```

ايلا درتي

```
reg.coef_
reg.intercept_
```

عا يخرج ليك α و β



يتبع..

الانحدار الخطي سهل للفهم و كايصلح لشحال من مشكل رياضياتي، و لكن را ماشي ديما هو الحل. مثلا ايلا شفتي البيانات اللي فالتصويرة، كايبان ليك ان انحدار خطي ماشي احسن اختيار فهاذ الحالة. فبحال هاذ الحالات كاندوزو للانحدار غير الخطي Nonlinear Regression اللي عالم كبير و متنوع...

شكرا على التتبع !

Noisy Data Sample
Nonlinear Regression

y

x

Source: [3]



المصادر و روابط للتعمق

1. صاوبت واحد دفتر جوبيتير اللي بغا يشوف واحد المثال ديال الانحدار الخطي و الكود ديالو

   Jupyter notebook

2. اللي بغا يتعمق كانصح بالدرس ديال الاستاذ الرائع Andrew Ng
   `https://fr.coursera.org/learn/machine-learning`
3. مصادر الصور

[1] http://pgfplots.net/tikz/examples/regression-line
[3] https://stackoverflow.com/questions/34474767/how-to-select-regression-algorithm-for-noisy-scattered-data/34476605

